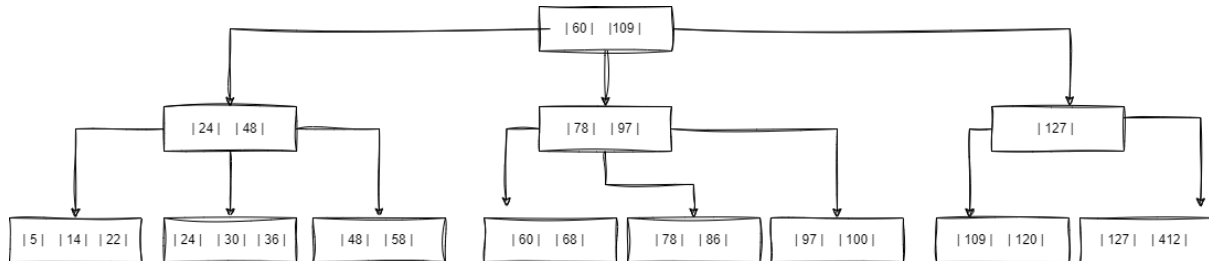


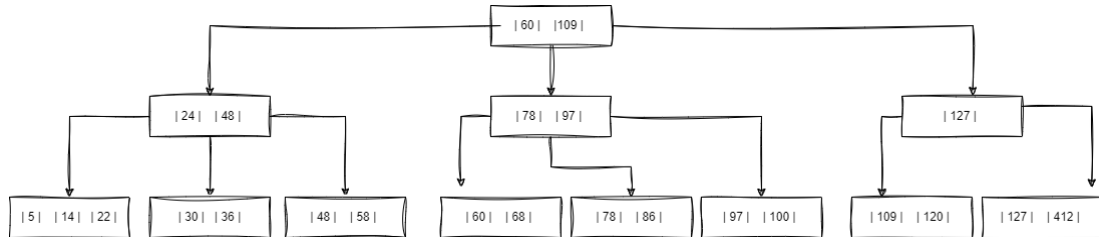
# מטלת מנחה 15 - קורס 20277

## שאלה 1

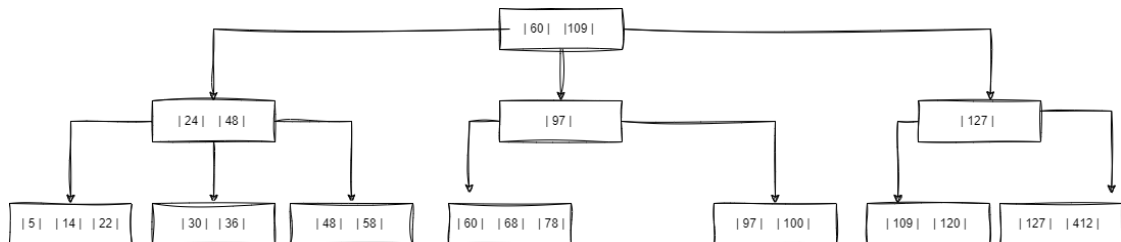
א. עץ ה-B+ שיווצר לאחר הכנסת כל הערכים:



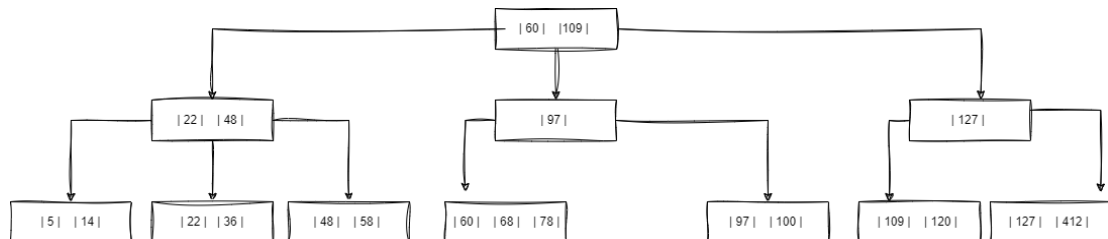
ב. לאחר מחיקת הערך 24:



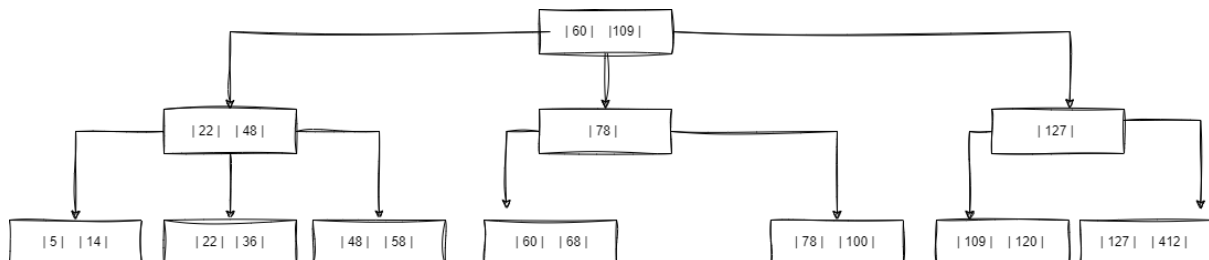
לאחר מחיקת הערך 86:



לאחר מחיקת הערך 30:



לאחר מחיקת הערך 100:



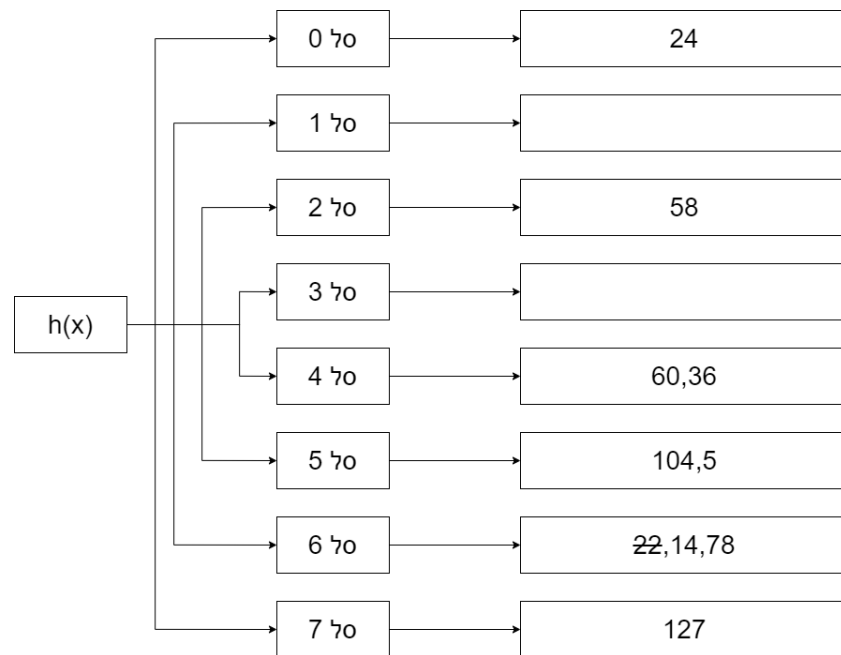
ג. העץ בסעיף א מכיל 3 רמות כך שהרמה הראשונה בעלת 2 ערכים. על מנת שהעץ יגדל בגובהו, יש להוסיף לרמה זו עוד שני ערכים (וכעת יהיו ברמה זו ח ערכים, ונידרש לעוד רמה).

ניתן להוסיף ערך נוסף לשורש בעזרת הוספת שני ערכים לבנו השמאלי. לבן זה יש 2 בנים בעלי 3 ערכים, ומספיקים **שני ערכים** להוסיף לבן זה שני ערכים, ולכן לשורש ערך נוסף. סך הכל נידרש להוסיף עוד **ארבעה ערכים** לפחות.

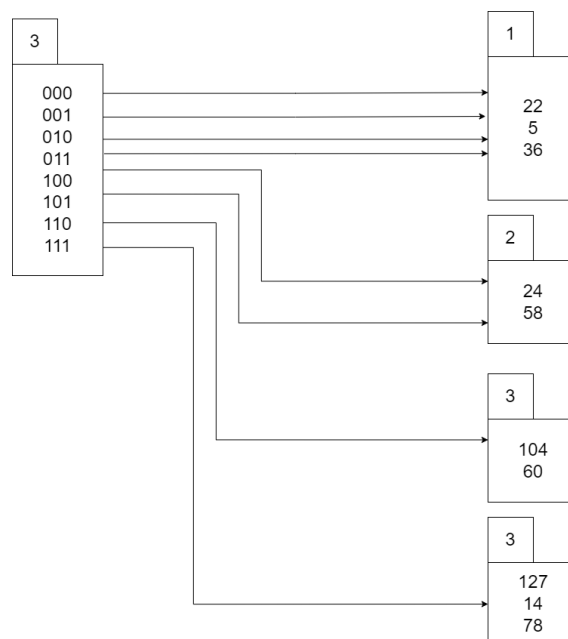
נתייחס לבן האמצעי של השורש. על מנת להוסיף לו 2 ערכים, יש להוסיף שני ערכים לכל אחד מבניו, סך הכל 4 ערכים לצאצאיו, ונקבל שיש להוסיף לפחות **שישה ערכים** על מנת להגביה את העץ.

## שאלה 2

א.



ב.



ג. במבנה של גיבוב מתרחב בעל רוחב תחילית גלובלית השווה ל  $n$ , לכל היותר  $2^{n-1}$  כניסות במדריך יכולות להצביע לסל אחד.

הסבר: מצב קיצון זה יתאפשר רק אם הסל מעולם לא "התפצל". תחילה, כאשר  $n = 1$ , הצביע אליו סל אחד. לאחר פיצול כלשהו,  $n = 2$  והצביעו אליו שני סלים. לאחר פיצול נוסף,  $n = 3$  והצביעו אליו 4 סלים (כמו בדוגמה בסעיף ב), וכך הלאה. עבור  $n = 11$  נקבל  $2^{10} = 1024$  סלים.

## שאלה 3

מפת הסיביות עבור העמודה region:

cid	11	22	33	44	55	66	77	88	99
north	0	1	0	1	0	0	0	0	0
south	0	0	0	0	1	1	1	0	0
hills	1	0	1	0	0	0	0	1	1

## שאלה 4

לאורך כל פתרון השאלה נשתמש בנוסחה מהמדריך לחישוב גודל יחס התוצאה  $p \times q$  על בסיס קבוצת התכונות A:  
 אם נסמן  $V = \max \{(A, p), V(A, p)\}$ , אז על פי עמוד 294 במדריך, מספר השורות ביחס התוצאה יהיה

$$\text{בערך } \frac{n_p \cdot n_q}{V}$$

א. אם נניח שערכי B מתפלגים באופן אחיד, אז השורות ביחס  $r$  המקיימים  $B = 85$  יהיה בערך  
 . כמו כן יש לכל היותר 40 ערכי C שונים ביחס זה, אולם  $V(C, s) = 20,000$  ולכן  
 נבחר ערך זה במונה.

מספר הערכים ביחס יהיה:

$$\frac{40 \cdot 30,000}{20,000} = 60$$

ב. שוב, על פי הנוסחה:

$$\frac{n_s \cdot n_q}{V(E, s)} = \frac{30,000 \cdot 10,000}{9,000} = 33,333$$

ג. יש לבצע חיתוך בין 15,000 ערכים שונים של A ל-1000 ערכים שונים של F.  
 נקבל לכל הפחות 0 ערכים משותפים ולכל היותר 1,000. אין סיבה להניח שתחום הערכים משותף, לכן סביר  
 להניח שמספר השורות בחיתוך יהיה קרוב יותר לאפס.

## שאלה 5

א. השאילתה מחזירה זוגות של עיר בצפון ומפלגה, כך שבמערכת בחירות כלשהי המפלגה קיבלה בעיר בין 1 ו-19 קולות, כולל.

ב. נשתמש בתכונת החילופיות והקיבוציות.  
ראשית נחשב את גודל תוצאת השאילתה  $city \times votes$ . על פי הנוסחה עליה הסברתי בשאלה 4, גודל היחס הוא:

$$\frac{150 \cdot 400,000}{150} = 400,000$$

אם נניח כי הערים מתפלגות באופן אחיד באזורים השונים בארץ, אז חמישית מהערים נמצאות באזור הצפון, ולכן יהיו 80,000 שורות של צירוף עיר בצפון-הצבעה.  
רק ב 2% מהמקרים הצביעו פחות מ-20 אנשים. גם כאן, אם נניח כי ההצבעות מתפלגות באופן אחיד ובלתי תלוי, נקבל כי סך הכל יהיו כ-1600 שורות בתוצאה.

ג. לכל היותר- בכל עיר הצביעו לכל מפלגה בכל מערכת בחירות. סך הכל נקבל:

$$200 \cdot 10 \cdot 150 = 30,000$$

המשמעות היא שבמקרה הגרוע מספר הנתונים דורש מקום אחסון רב, וייתכן שכדאי לשקול להשתמש באינדקס על מנת לגשת אליו בקלות.

לכל הפחות - היות ובכל מערכת בחירות היו מצביעים שהצביעו בה, במקרה הגרוע נקבל רשומה של הצבעה מעיר אחת עבור כל מערכת בחירות, כלומר 20 שורות.

אם ניקח בחשבון את הנתון  $V(cid, votes) = 150$ , אז מספר השורות המינימלי יהיה 150.